# Data Management for Climate Science

## Lessons Learned and Future Needs

**Ilana Stern, NCAR**

*ilana@ucar.edu*

# Outline

- NCAR HPC and big data
- CMIP3 and CMIP5
- Evolution of CESM workflow
- Future plans and needs

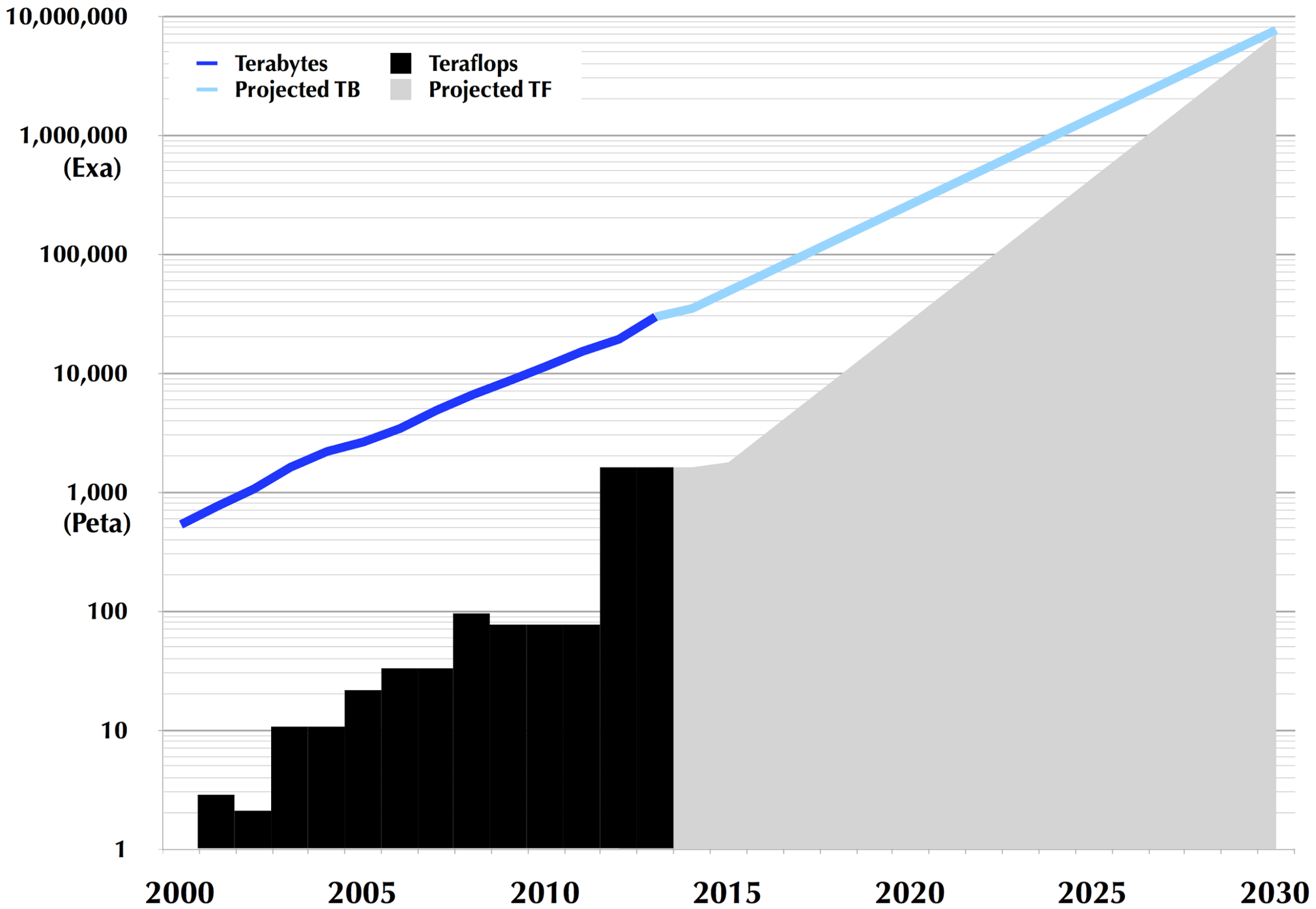

Ilana Stern ilana@ucar.edu

# flops vs. bytes, year reached @ NCAR

| SI prefix | flops | bytes |
|---|---|---|
| giga- | 1988 | long ago |
| tera- | 2000 | < 1986 |
| peta- | 2012 | 2002 |
| exa- | 2024? | 2025 |
| zetta- | 2036? | ? |


Ilana Stern ilana@ucar.edu

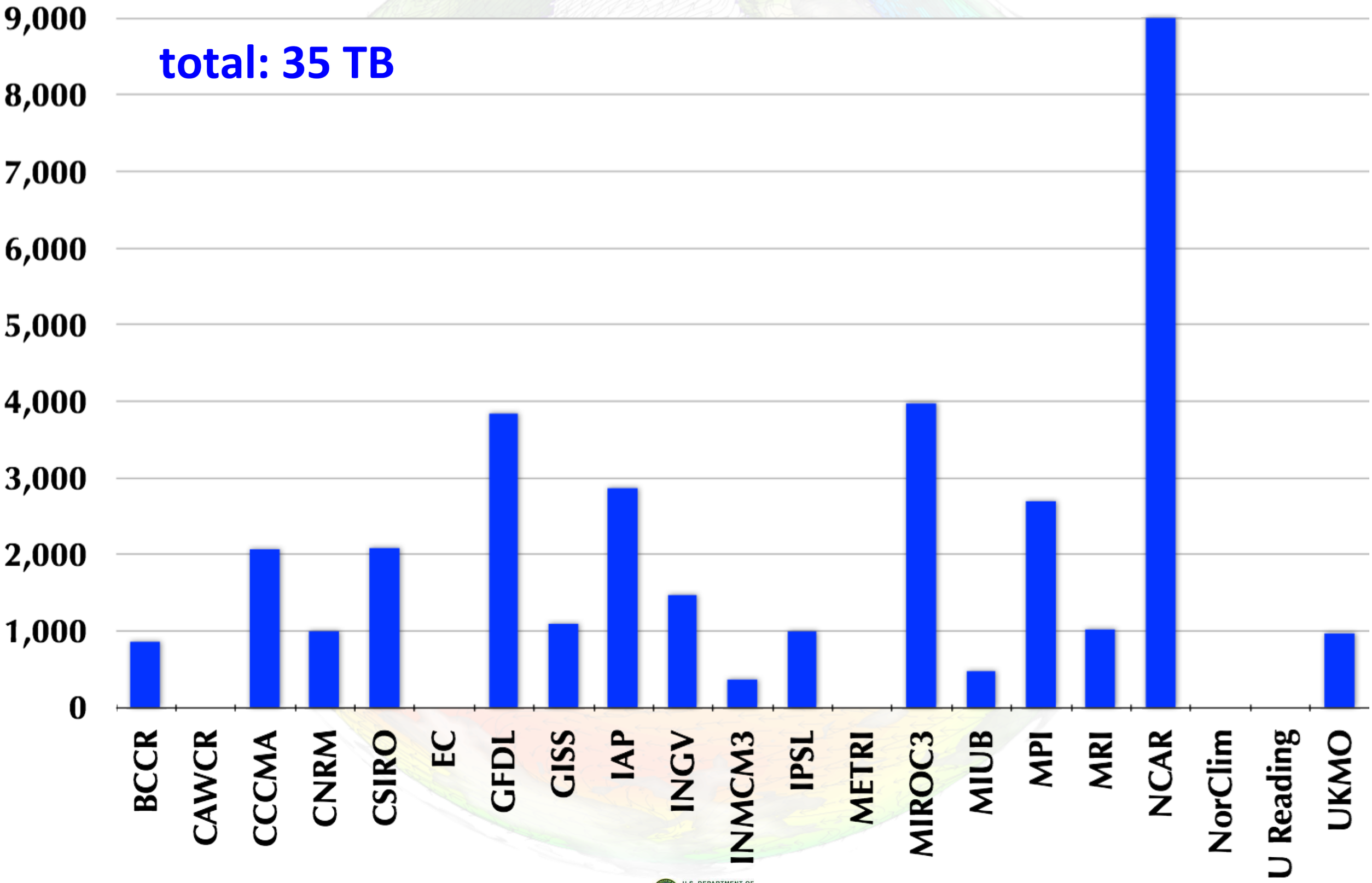
# NCAR flops and bytes, 2000-2030





Ilana Stern ilana@ucar.edu

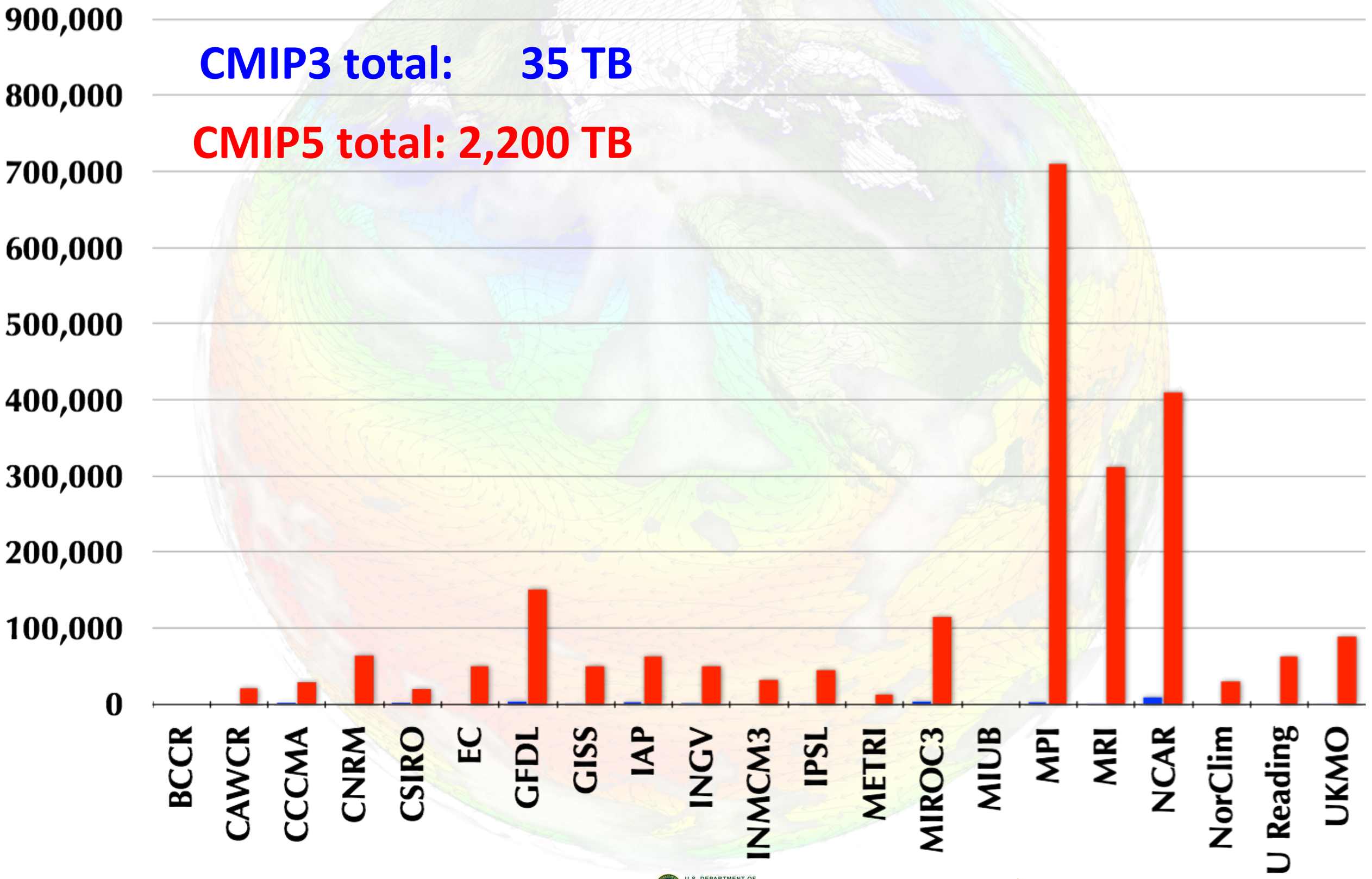
# CMIP3/AR4 data volumes by group





Ilana Stern ilana@ucar.edu

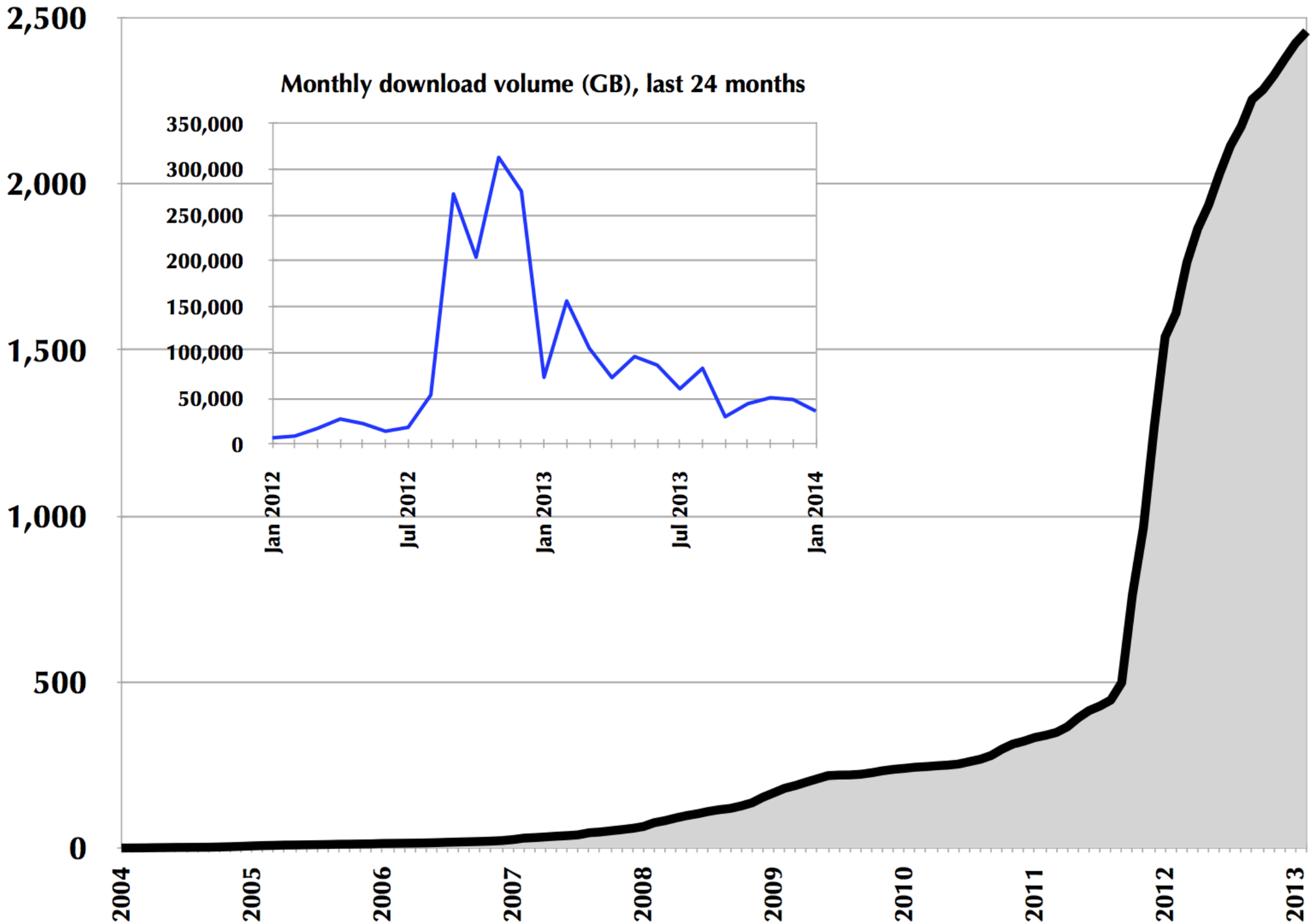
# CMIP5/AR5 data volumes by group





Ilana Stern ilana@ucar.edu

# NCAR ESG-CET Portal Downloads (TB)





Ilana Stern ilana@ucar.edu

# Workflow prior to late 1990s

## model

| time 1 | time 2 | ... | time m |
|---|---|---|---|
| header | header | | header |
| field 1 | field 1 | | field 1 |
| field 2 | field 2 | | field 2 |
| ... | ... | | ... |
| field n | field n | | field n |

## small disk

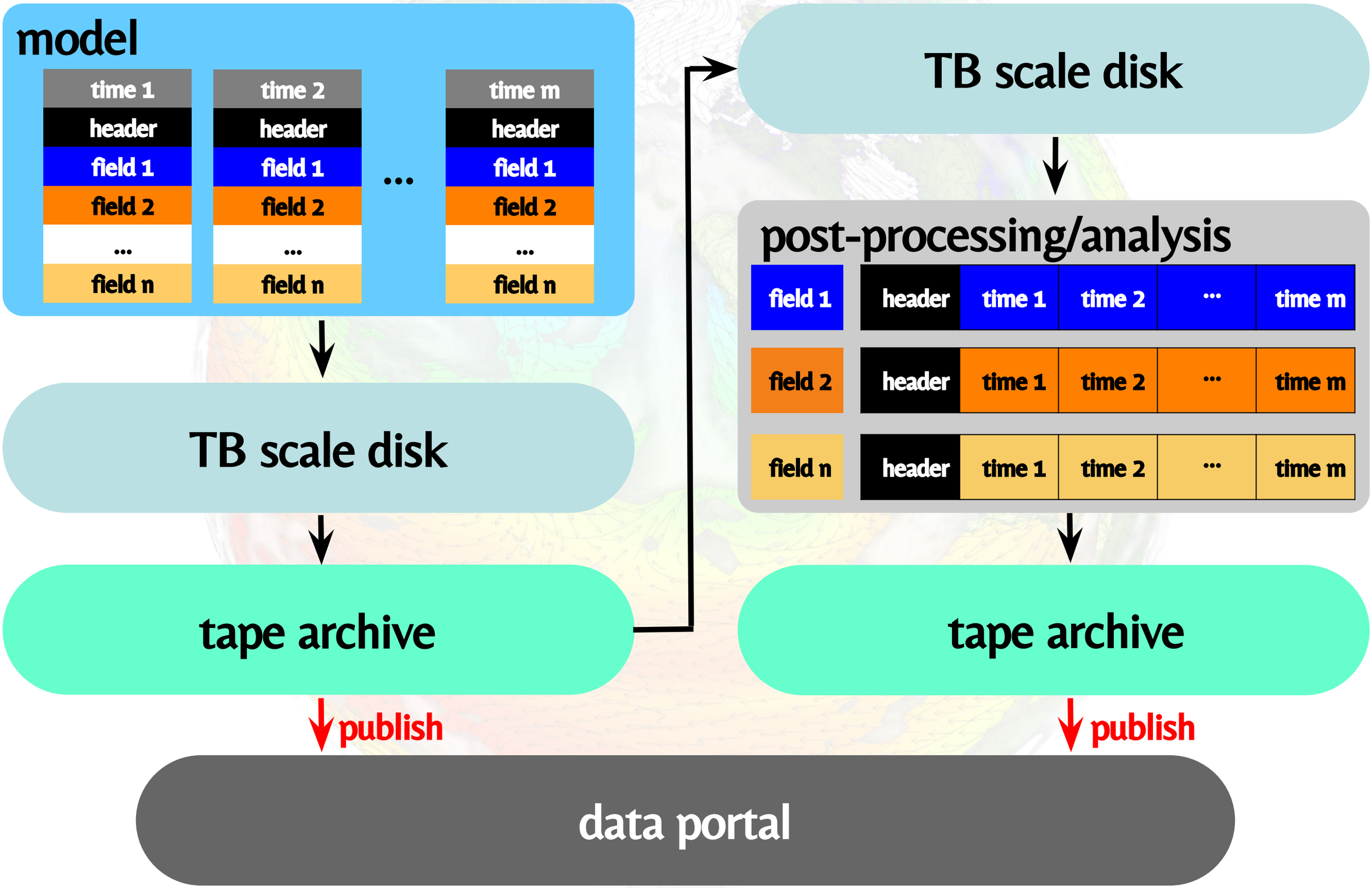
## tape archive

## later

## post-processing/analysis


Ilana Stern ilana@ucar.edu

# Workflow 2000-2012





Ilana Stern ilana@ucar.edu

# Workflow 2000-2012

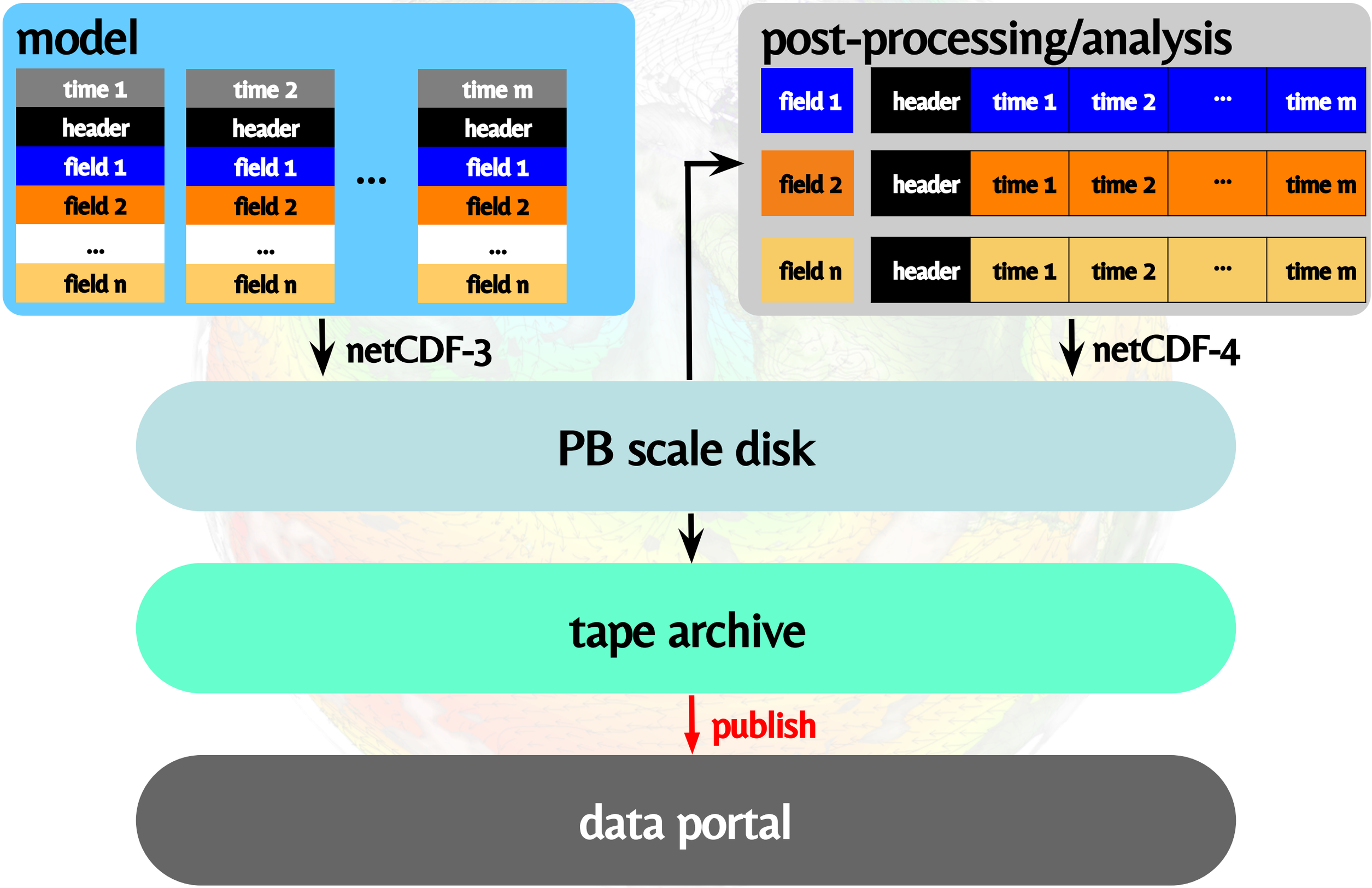
## Typical workflow for standard runs, CMIPs

- Model writes time-slice (aka “history”) files from disk to tape
- Files read back from tape to disk for postprocessing
- Post-processed files (aka “time-series”) written to tape
- If for MIP, convert time-series files with CMOR
- If necessary for MIP, much easier to work with time-series data

## Issues:

- Many writes/reads to/from tape
- Multiple copies of identical data
- Constrained by “small” disk space
- Hugely inefficient

# Current workflow





Ilana Stern ilana@ucar.edu

# Current workflow

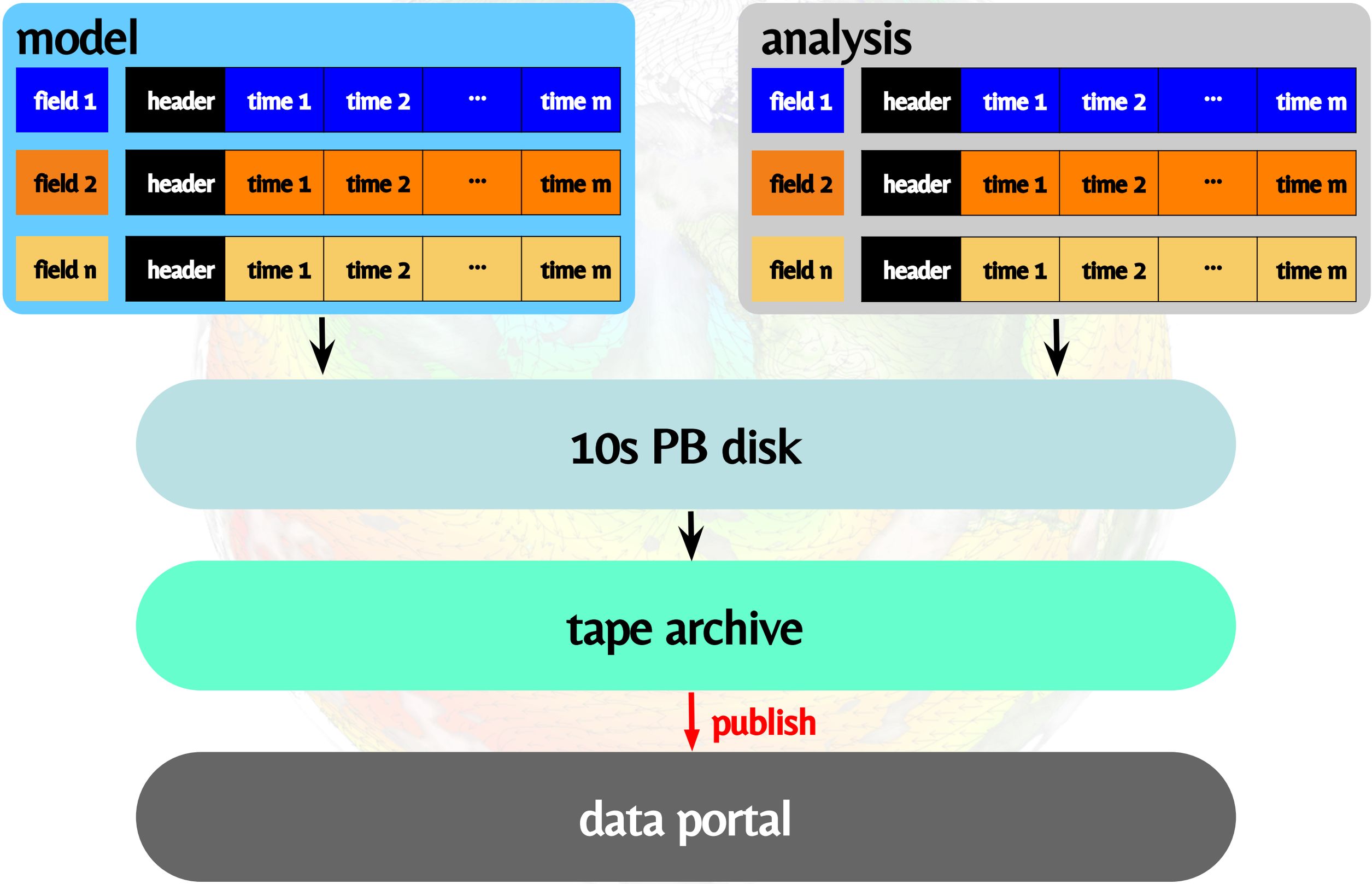
- Model writes time-slice files to disk **only**
- Files postprocessed *in situ*
- **Only** time-series files written to HPSS
- Currently uses separate serial (nc3) postprocessing scripts
  - Within 3 months: parallel (nc4) postprocessing scripts
  - Eventually: PIO (Parallel I/O) package to be rewritten to write single-field format directly from model

## Advantages

- Many fewer writes to HPSS - no reads
- Possible because of sufficient disk space now available
- Single copies of data
- Much more efficient for analysis purposes


Ilana Stern ilana@ucar.edu

# Near-term future Workflow





U.S. DEPARTMENT OF ENERGY Office of Science NCAR NSF Ilana Stern ilana@ucar.edu

# Near-term future workflow

I/O package writes single-field format directly from model

## Advantages

- Many fewer writes to HPSS - no reads
- Possible because of sufficient disk space now available
- Single copies of data
- Much more efficient
- Potential MIPs are potentially written directly from model, no additional steps required

## Issues

- Potential difficulties with reruns or branches
- Possible to create unwieldy very large (1 TB+) files
- Granularity of file sizes very dependent on exact configuration

# Exascale workflow

modeling

analysis

analyzed/derived data

100s PB disk

publish

tape archive (?)

publish (?)

data portal (analyzed/derived data only)


Ilana Stern ilana@ucar.edu

# Current CESM big data projects

## Current

<u>CESM1-CAM5-BGC ensemble</u>

- 40 runs, total ~7,500 model years, ~400 TB

<u>Last millennium ensemble</u>

- 26 runs, total ~26,000 model years, ~600 TB

(Both using newest workflow)

## <u>Longer-term big data</u>

CMIP6 (2016-2017)

Potential additional -MIPs

Higher resolution (1/8° SE atm/lnd, 1/10° ocn/ice)


Ilana Stern ilana@ucar.edu

# Current CESM big data projects

| Category | CMIP5 | Large Ensemble | Last Millennium Ensemble |
|---|---|---|---|
| **Model(s)** | CCSM4<br>CESM1-CAM5<br>CESM1-BGC<br>CESM1-WACCM<br>CESM1-FASTCHEM | CESM1-CAM5-BGC | CESM1-CAM5 |
| **Volume measure 1** | ~1,600 TB | ~750 TB | ~1,000 TB |
| **Volume measure 2** | ~180 TB | ~300 TB | ~420 TB |
| **Total simulated years** | ~30,000 | ~7,500 | ~26,000 |
| **Number of runs** | 555 | 62 (+ at least 10 more) | 26 |
| **Output categories** | 19 | 14 | 7 |
| **Number of fields** | 951 | 1127 | 820 |


Ilana Stern ilana@ucar.edu

# DOIs for data

- UCAR-wide and institutionally-supported effort to add Digital Object Identifiers (DOIs) to UCAR data holdings
  - Effort led by the NCAR Library with participation by the Research Data Archive (RDA), EOL, CESM, HAO, etc.
  - NCL and Yellowstone have been assigned DOIs, among a number of RDA-held datasets and EOL-held datasets
- Intent is to make datasets citable for "forever", enabling reproducibility and traceability for UCAR science
- At this time, does not imply data are "peer-reviewed"
- A "permanent" pointer to a landing page for data
- Give credit to data creators and data managers and scientists for their work
- UCAR uses the EZID service to generate random DOIs with the assigned-to-UCAR prefix http://dx.doi.org/10.5065
- DOIs are stored by DataCite and the responsible unit maintains them


Ilana Stern ilana@ucar.edu

# CESM and DOIs

- CESM-CAM5-BGC Large Ensemble and Last Millennium Ensemble to serve as test cases
- Issues include “granularity” - at what level (experiment, component, component stream, individual file) are DOIs assigned?
- What about simulations done elsewhere (DOE, etc.)?
- Permanence (really?) and responsibility
- Coordination of DOIs with MIP projects - how do we link CMIP5 data to the original CESM output, once the latter has a DOI?


Ilana Stern ilana@ucar.edu

# CESM plans and future needs

## Issues

- User community needs/wants drives all!
- Modeling and analysis ~concurrently to avoid memory -> disk latency and all the other issues
- Ongoing updates of workflow
- Updates to CESM data management policy to reflect workflow and other changes
- Longer-term viability of ESG/ESGF model - downloading PB isn't sustainable (?)
- Must have serious server-side analysis
- Possibility of reruns instead of save everything


Ilana Stern ilana@ucar.edu

# Thank you!

# Questions?


Ilana Stern ilana@ucar.edu